#6

```
                    RAW SEQUENCE LISTING          DATE: 10/20/2000
                    PATENT APPLICATION:  US/09/557,796     TIME: 11:48:15

                    Input Set : A:\557796.txt
                    Output Set: N:\CRF3\10202000\I557796.raw


     4 <110> APPLICANT: Hoch, James
     5        Dartois, Veronique
     7 <120> TITLE OF INVENTION: METABOLIC SELECTION METHODS
    10 <130> FILE REFERENCE: 234/191
    12 <140> CURRENT APPLICATION NUMBER: 09/557,796
    13 <141> CURRENT FILING DATE: 2000-04-25
    15 <150> PRIOR APPLICATION NUMBER: 09/172,952
    16 <151> PRIOR FILING DATE: 1998-10-14
    18 <160> NUMBER OF SEQ ID NOS: 33
    20 <170> SOFTWARE: FastSEQ for Windows Version 3.0
    22 <210> SEQ ID NO: 1
    23 <211> LENGTH: 816
    24 <212> TYPE: DNA
    25 <213> ORGANISM: yia j
    27 <400> SEQUENCE: 1
    29  atgggcacaa aagaaagcga gaacacgcaa gataaagaga ggcctgccgg aagtcagagc       60
    30  ctttttcgtg ggttgatgct aattgagatc ctgagtaatt atccaaatgg ctgtcccgtg      120
    31  gcgcatctgt cggaactggc gggactgaac aaaagtaccg ttcatcgctt attacagggg      180
    32  ctgcagtcct gcgggtacgt gacgcctgcc ccggcggcgg ggagctatgc gctgacgaca      240
    33  aaatttatcc gcgttggcca aaaggcgttg tcgtcgctga atattatcca cgtcgcggcg      300
    34  ccgcatcttg aggcgcttaa cctgccacc ggcgagacgg tgaacttctc cagccgtgaa       360
    35  gatgaccacg cgatcctgat ttataagctg gagccgacca ccggtatgct gcgtacgcgc      420
    36  gcctatattg gccagcacat gcgctgtact gctcggcaat gggcaaagat ttatatggcg      480
    37  tttggccatc ctgactacgt tgagagctac tggaattcac accaggagat tatccagccg      540
    38  ctgacccgta ataccattac cggcttgcct gcgatgcatg atgaactggc gcagatccgc      600
    39  gagcgaaata tggcgatgga cagggaagag aacgagctgg gcgtgtcgtg cctggctgtc      660
    40  cccgtttttg atatccatgg gcgcgtgcct tatgccattt ctatctctct atcaacatcg      720
    41  cgcctcaagc aggtgggaga gaaaaattta ctcaagccgc tacgcgatac ggcagaggcg      780
    42  atttctcgcg aactgggctt ttccgtgcgg gaaggt                               816
    45 <210> SEQ ID NO: 2
    46 <211> LENGTH: 996
    47 <212> TYPE: DNA
    48 <213> ORGANISM: yia k
    50 <400> SEQUENCE: 2
    52  atgaaagtca cgtttgagca gttaaaagag gcattcaatc gggtactgct ggacgcgtgc       60
    53  gtcgcccggg aaaccgccga tgcctgcgca gaaatgtttg cccgcaccac cgaatccggc      120
    54  gtctattctc acggcgtgaa ccgctttcct cgcttcatcc agcagttgga taacggcgac      180
    55  attatccctg aggctcaacc gcagcgggtg accacgctcg gcgccatcga acagtgggat      240
    56  gctcagcgtt ccatcggcaa cctgacggcg aaaaagatga tggatcgggc cattgagctg      300
    57  gcctccgatc acggtatcgg cctggtcgcc ttacgtaatg ctaaccactg gatgcgcggc      360
    58  ggcagctacg gctggcaggc ggcggaaaaa ggctacatcg gtatctgctg gaccaactcc      420
    59  atcgccgtta tggcgccatg gggcgctaaa gagtgccgta tcggtaccaa cccgctgatc      480
    60  gtcgccattc cgtcgacgcc gatcaccatg gtggatatgt cgatgtcgat gttctcctac      540
    61  ggcatgctgg aggttaaccg ccttgccggc cgcgaactgc ccgtggacgg cggattcgac      600
    62  gatgacggtc gtttgaccaa agagccgggg acgatcgaga aaaatcgccg catttttaccc      660
    63  atgggctact ggaaaggttc cggcctgtcg atcgtgctgg atatgattgc caccctcctc      720
    64  tccaacggat cgtcggttgc cgaagtgacc caggaaaaca gcgatgaata tggcgtttcg      780
```

```
        RAW SEQUENCE LISTING              DATE: 10/20/2000
        PATENT APPLICATION:  US/09/557,796   TIME: 11:48:15

        Input Set : A:\557796.txt
        Output Set: N:\CRF3\10202000\I557796.raw
```

```
65  cagatcttca tcgctattga agtggataag ctgatcgacg gcgcaacccg cgacgccaag    840
66  ctgcaacgga ttatggattt catcaccacc gccgagcgcg ccgatgaaaa tgtggcggtc    900
67  cgtcttcctg gccatgaatt tacccgtctg ctggatgaaa accgccgcaa cggcattacc    960
68  gtcgatgaca gcgtatgggc caaaattcag gcgctg                             996
71  <210> SEQ ID NO: 3
72  <211> LENGTH: 462
73  <212> TYPE: DNA
74  <213> ORGANISM: yia 1
76  <400> SEQUENCE: 3
78  atgatttttg gtcatattgc tcaacctaat ccgtgtcgtc tgcccgcggc cattgagcgg     60
79  gcgcttgatt tcctgcgcac gacggatttc cacgcgctgg cacccggcgt cgtggaaatc    120
80  gacggccaaa acatcttcgc gcaggttatc gacttaacca ctcgcgatgc cgctgaaaat    180
81  cgtccggagg tccaccgtcg ctatctggat atccagtttc tggcatcggg cgaagaaaaa    240
82  atcggtatcg ccattgatac cggcaataat caaatcagcg aatctttatt agaacagcgc    300
83  gatattattt tttatcacga cagcgaacat gaatcgttct ttgaaatgac gccaggcaac    360
84  tatgcgatat ttttcccgca agatgttcat cgtcctgtat gtaataaaac tgtagccacg    420
85  ccgatccgca aaatagtcgt taaagtcgct atttcagttt ta                      462
88  <210> SEQ ID NO: 4
89  <211> LENGTH: 945
90  <212> TYPE: DNA
91  <213> ORGANISM: orf1
93  <400> SEQUENCE: 4
95  atgaattcga ataataccgg ttacattatc ggtgcgtacc cctgtgcccc ctgtgcaccc     60
96  tcatttcacc aaaagagtga agaggaagag atggaattct ggcggcagct ctccgacacc    120
97  ccggatattc gcgggctgga gcaaccctgc ctaccctgcc ttgaacatct tcatccgctc    180
98  ggcgacgagt ggttattgcg ccataccccg ggacactggc agattgtcgt taccgccatc    240
99  atggaaacca tgcgccgccg cggtgaaaac ggcggctttg ggctggcgtc cagcgacgaa    300
100 acgcagcgca aagcctgcgt ggagtactat cgccacctgc agcagaagat cgctaaaatc    360
101 aatggcaata ccgccggaaa ggtcattgcc cttgagcttc acgccgcccc gctggcgggc    420
102 aatgccaacg tggctcaggc taccgacgcc tttgcccgtt cattaaaaga aattacccgc    480
103 tgggactggt cctgcgagct ggtgctggag cactgcgacg cgatgaccgg cagcgcgccg    540
104 cgcaaaggat ttttgccgtt agaaaacgtg ctggaagcca ttgccgatta tgacgttggc    600
105 atttgtatta actgggcgcg ttcggccatt gaagggcgga ataccgtgct accgctcacc    660
106 catacgcagc aggtaaaacg ggcaggaaag ctcggcgacg tgatgttttc tggcacgacg    720
107 cagaccggcg agtacggcga atggcaggat ttacacgcgc cgttcgcgcc tttctgcccg    780
108 cagagcctga tgaccaccga acacgctcgt gaattatttg cctgcgcagg aaccgccccc    840
109 ctgcaatttt caggcattaa attactggaa attaatgcca gcgcaaacgt tgatcatcgc    900
110 atcgcgatat tacgcgacgg catctccgcg ctaaaacaag cacaa                   945
113 <210> SEQ ID NO: 5
114 <211> LENGTH: 1317
115 <212> TYPE: DNA
116 <213> ORGANISM: yia x2
118 <400> SEQUENCE: 5
120 atgaatataa cctctaactc tacaaccaaa gatataccgc gccagcgctg gttaagaatc     60
121 attccgccta tactgatcac ttgtattatt tcttatatgg accgggtcaa tattgccttt    120
122 gcgatgcccg gaggtatgga tgccgactta ggtatttccg ccaccatggc ggggctggcg    180
123 ggcggtattt tctttatcgg ttatctattt ttacaggttc ccggcgggaa aattgccgtt    240
124 cacggtagcg gtaagaaatt tatcggctgg tcgctggtcg cctgggcggt catctccgtg    300
125 ctgacggggt taattaccaa tcagtaccag ctgctggccc tgcgcttctt actgggcgtg    360
```

```
126  gcggaaggcg gtatgctgcc ggtcgttctc acgatgatca gtaactggtt ccccgacgct   420
127  gaacgcggtc gcgccaacgc gattgtcatt atgtttgtgc cgattgccgg gattatcacc   480
128  gccccactct caggctggat tatcacggtt ctcgactggc gctggctgtt tattatcgaa   540
129  ggtttgctct cgctggttgt tctggttctg tgggcataca ccatctatga ccgtccgcag   600
130  gaagcgcgct ggatttccga agcagagaag cgctatctgg tcgagacgct ggccgcggag   660
131  caaaaagcca ttgccggcac cgaggtgaaa aacgcctctc tgagcgccgt tctctccgac   720
132  aaaaccatgt ggcagcttat cgccctgaac ttcttctacc agaccggcat ttacggctac   780
133  accctgtggc tacccaccat tctgaaagaa ttgacccata gcagcatggg gcaggtcggc   840
134  atgcttgcca ttctgccgta cgtcggcgcc attgctggga tgttcctgtt ttcctccctt   900
135  tcagaccgaa ccggtaaacg caagctgttc gtctgcctgc cgctgattgg cttcgctctg   960
136  tgcatgttcc tgtcggtggc gctgaaaaac caaatttggc tctcctatgc cgcgctggtc  1020
137  ggctgcggat tcttcctgca atcggcggct ggcgtgttct ggaccatccc ggcacgtctg  1080
138  ttcagcgcgg aaatggcggg cggcgcgcgc ggggttatca acgcgcttgg caacctcggc  1140
139  ggattttgtg gcccttatgc ggtcgggggtg ctgatcacgt tgtacagcaa agacgctggc  1200
140  gtctattgcc tggcgatctc cctggcgctg gccgcgctga tggcgctgct gctgccggcg  1260
141  aaatgcgatg ccggtgctgc gccggtaaag acgataaatc cacataaacg cactgcg     1317
144  <210> SEQ ID NO: 6
145  <211> LENGTH: 1503
146  <212> TYPE: DNA
147  <213> ORGANISM: lyxk
149  <400> SEQUENCE: 6
151  atgagcaaga aacaggcctt ctggctgggt attgattgcg gcggcaccta tctgaaagcc    60
152  ggtttatatg acgccgaagg tcatgaacat ggcattgtgc ggcaagcgct acggacgatg   120
153  tcgcccctgc cgggttacgc cgaacgcgac atgcgccagc tctggcaaca ctgcgcggcg   180
154  accattgccg ggctattaca gcaggcaggt gtatccgtcg aacagattaa aggcgtgggc   240
155  atctccgctc agggtcaagg gctctttctc ctcgataagc aggatcggcc gctgggtaac   300
156  gccatcctct cctccgatcg tcgggcgctg aaaatcgttc agcgctggca gcgggaccgt   360
157  attcccgaac ggctctatcc cgttacccgc cagacgctgt ggaccggaca tccggcttct   420
158  ttgctgcgct gggtaaaaga gaatgaaccc cagcgctacg cgcaaattgg ctgcgtgatg   480
159  atgggggcatg actatctgcg ctggtgctta accggcgcga aggctgcga ggagagcaac   540
160  atctccgagt ccaacctcta caacatggcc atgggccagt acgacccgcg cctgaccgag   600
161  tggctgggca tcggtgaaat cgatagcgcg ctgcccccg ttgtagggtc agccgaaatt   660
162  tgcgggggaga tcaccgctca ggcagccgct ttaaccggtc tggcggcggg tactcccgtc   720
163  gttggcgacc tgtttgacgt ggtctccacc gccctttgcg ccgggattga ggatgagtcg   780
164  accctcaatg cggtgatggg gacctgggcc gtcactagcg gtatcgctca cggcctgcgc   840
165  gaccatgagg cccacccttta cgtctatggc cgctacgtca atgacggcca gtatatcgtt   900
166  cacgaagcca gcccgacctc atccggcaac ctcgaatggt ttaccgccca gtggggcgat   960
167  ctctcgtttg atgagatcaa tcaggccgtc gccagcctgc cgaaagccgg gagcgagctg  1020
168  ttttttctgc cgtttctgta tggcagcaac gccgggctgg agatgacctg cggcttttac  1080
169  ggcatgcagg cgctgcatac ccgcgcgcac ctgctgcagg cggtttatga aggcgtggta  1140
170  tttagccata tgacccaacct cagccgtatg cgcgaacgct ttacaaacgt tcaggccctg  1200
171  cgcgtcaccg cgggcccggc gcactccgac gtctggatgc agatgctggc ggacgtaagc  1260
172  ggcttacgca ttgaactccc gaaggtggaa gagaccgct gttttggcgc ggccctcgcc  1320
173  gctcgtgtcg gtaccggcgt ataccgcagc tttagcgaag cccggcgcgc ccggcagcac  1380
174  ccggtgcgca cgctgctgcc cgatatgacc gcccacgcgc gctatcagcg caaataccgc  1440
175  cactacctgc atttgattga agcactacag ggctatcacg cccgtattaa ggagcacgca  1500
176  tta                                                                1503
179  <210> SEQ ID NO: 7
180  <211> LENGTH: 660
```

RAW SEQUENCE LISTING　　　　　　　　DATE: 10/20/2000
PATENT APPLICATION:　US/09/557,796　　TIME: 11:48:15

Input Set : A:\557796.txt
Output Set: N:\CRF3\10202000\I557796.raw

```
181 <212> TYPE: DNA
182 <213> ORGANISM: yia q
184 <400> SEQUENCE: 7
186  atgagccgac cattactgca gctggcgctc gaccatacca gccttcaggc tgcgcagcgc   ·60
187  gatgtcgccc tgctacagga tcacgttgat attgtggagg cgggaaccat cctctgctta   120
188  accgaagggc ttagcgcggt taaagccctg cgcgcccagt gtccggggaa gatcatcgtc   180
189  gccgactgga aagtcgccga cgccggtgaa accctggcgc agcaggcctt tggcgctggc   240
190  gccaactgga tgaccatcat ttgcgccgca ccgctcgcca cggtcgagaa aggccacgcc   300
191  gtggcccagg cctgcggcgg tgaaattcag atggagctgt tcggcaactg gacgctggat   360
192  gacgcccgcg cctggtaccg taccggcgtc catcaggcga tttaccatcg cggacgcgat   420
193  gcccaggcca gcgggcagca gtggggggag gcggatctgg cgcgcatgaa agcgctgtcc   480
194  gatattggcc ttgagctatc gattaccggc ggcattaccc cagccgatct accgctgttc   540
195  aaagatatca acgtcaaagc ctttattgcc gggcgcgcgc tggcaggcgc cgcccatccg   600
196  gcgcggggtg ccgccgaatt ccacgcgcaa atcgacgcta tctggggaga acagcatgcg   660
199 <210> SEQ ID NO: 8
200 <211> LENGTH: 858
201 <212> TYPE: DNA
202 <213> ORGANISM: yia r
204 <400> SEQUENCE: 8
206  atgcgtaacc acccgttagg tatttatgaa aaagcgctgg cgaaggatct cagctggcct   60
207  gagcggctgg tactggccaa aagctgcggt tttgattttg. tcgaaatgtc ggtggacgag   120
208  accgatgaac gcctttcgcg cctggagtgg accccggccc agcgcgcatc gctggtgagc   180
209  gcgatgctgg aaaccgcggt cgccattccc tcgatgtgct tgtccgccca tcgccgtttc   240
210  ccctttggca gccgcgatga agcggtacgc gatcgggcgc· gagagattat gaccaaagcc   300
211  atccgcctgg cgcgcgatct ggggatccgc accatccagc tggcgggta cgacgtctat   360
212  tacgaagagc atgatgaagg cacccggcag cgttttgccg aagggctggc ctgggcggta   420
213  gaacaggccg ccgccgcgca ggtaatgctg gcggtggaga tcatggacac cgcctttatg   480
214  aactccatca gcaaatggaa aaagtgggac gagatgcttt cgtcaccgtg gtttaccgtc   540
215  tacccggacg tcggcaacct cagcgcctgg ggaaacgacg tcaccgccga gctgaagctg   600
216  ggcatcgatc gtatcgccgc catccacctg aaagatacgc tgcccgtgac cgacgatagc   660
217  cctggccagt tccgcgacgt gccgttcggc gaaggatgcg tcgatttttg tcggcattttt   720
218  aagacgctgc gcgagctgaa ctaccgcggt tcattttttga ttgagatgtg gacggagaaa   780
219  gccagcgagc cggtgctgga gattatccag gcccggcgct ggatcgaatc acggatgcag   840
220  gaaggggggat tcacatgt                                                858
223 <210> SEQ ID NO: 9
224 <211> LENGTH: 714
225 <212> TYPE: DNA
226 <213> ORGANISM: yia s
228 <400> SEQUENCE: 9
230  atgttagaac aactgaaagc cgaggtactg gcggcaaacc tggccctccc cgcacacggc   60
231  ctggtcacct ttacctgggg caacgtcagc gcggtcgatg aaacgcgcaa gctgatggtc   120
232  attaagcctt ccggcgtcga atatgaggtg atgaccgccg acgatatggt ggtcgtagag   180
233  atggccagcg gtaaagtcgt tgaaggcggt aaaaaaccct cttcagatac gccaacgcat   240
234  ctggcgcttt atcgcgcgcta tccgcagatc ggcgggatcg tgcatacccca ctcccgccac   300
235  gcgacgatct ggtcgcaggc cgggctcgat ctccccgcct ggggcaccac ccacgccgac   360
236  tacttctatg gcgcgatccc ctgtacccga cggatgaccg ttgaggagat taacggcgag ·  420
237  tatgagtatc agaccggcga ggtgattatc aaaacctttg aacagcgcgg cctggatccg   480·
238  gcgcaaatcc cggcggtatt ggtccattca cacgcgcccct ttgcctgggg taaagacgcc   540
239  gccgacgccg tacataacgc cgtggtgctg gaggagtgcg cctacatggg cctcttctcg   600
```

```
240  cgccagtggc cacagctgcc ggatatgcag tctgaactgc tcgataaaca ctatctgcgt    660
241  aaacacggcg cgaacgctat tacgggcaaa actagtcccg cggaactccc cgga          714
244  <210> SEQ ID NO: 10
245  <211> LENGTH: 272
246  <212> TYPE: PRT
247  <213> ORGANISM: YiaJ-Ko
249  <400> SEQUENCE: 10
251  Met Gly Thr Lys Glu Ser Glu Asn Thr Gln Asp Lys Glu Arg Pro Ala
252  1               5                   10                  15
254  Gly Ser Gln Ser Leu Phe Arg Gly Leu Met Leu Ile Glu Ile Leu Ser
255              20                  25                  30
257  Asn Tyr Pro Asn Gly Cys Pro Val Ala His Leu Ser Glu Leu Ala Gly
258          35                  40                  45
260  Leu Asn Lys Ser Thr Val His Arg Leu Leu Gln Gly Leu Gln Ser Cys
261      50                  55                  60
263  Gly Tyr Val Thr Pro Ala Pro Ala Ala Gly Ser Tyr Ala Leu Thr Thr
264  65                  70                  75                  80
266  Lys Phe Ile Arg Val Gly Gln Lys Ala Leu Ser Ser Leu Asn Ile Ile
267                  85                  90                  95
269  His Val Ala Ala Pro His Leu Glu Ala Leu Asn Leu Ala Thr Gly Glu
270                  100                 105                 110
272  Thr Val Asn Phe Ser Ser Arg Glu Asp Asp His Ala Ile Leu Ile Tyr
273              115                 120                 125
275  Lys Leu Glu Pro Thr Thr Gly Met Leu Arg Thr Arg Ala Tyr Ile Gly
276          130                 135                 140
278  Gln His Met Arg Cys Thr Ala Arg Gln Trp Ala Lys Ile Tyr Met Ala
279  145                 150                 155                 160
281  Phe Gly His Pro Asp Tyr Val Glu Ser Tyr Trp Asn Ser His Gln Glu
282                  165                 170                 175
284  Ile Ile Gln Pro Leu Thr Arg Asn Thr Ile Thr Gly Leu Pro Ala Met
285              180                 185                 190
287  His Asp Glu Leu Ala Gln Ile Arg Glu Arg Asn Met Ala Met Asp Arg
288          195                 200                 205
289  Glu Glu Asn Glu Leu Gly Val Ser Cys Leu Ala Val Pro Val Phe Asp
290      210                 215                 220
292  Ile His Gly Arg Val Pro Tyr Ala Ile Ser Ile Ser Leu Ser Thr Ser
293  225                 230                 235                 240
295  Arg Leu Lys Gln Val Gly Glu Lys Asn Leu Leu Lys Pro Leu Arg Asp
296                  245                 250                 255
298  Thr Ala Glu Ala Ile Ser Arg Glu Leu Gly Phe Ser Val Arg Glu Gly
299                  260                 265                 270
303  <210> SEQ ID NO: 11
304  <211> LENGTH: 332
305  <212> TYPE: PRT
306  <213> ORGANISM: YiaK-Ko
308  <400> SEQUENCE: 11
310  Met Lys Val Thr Phe Glu Gln Leu Lys Glu Ala Phe Asn Arg Val Leu
311  1               5                   10                  15
313  Leu Asp Ala Cys Val Ala Arg Glu Thr Ala Asp Ala Cys Ala Glu Met
```

VERIFICATION SUMMARY   DATE: 10/20/2000
PATENT APPLICATION: US/09/557,796   TIME: 11:48:16

Input Set : A:\557796.txt
Output Set: N:\CRF3\10202000\I557796.raw